DOCUMENT RESUME

ED 030 687 UD 007 269

By-Carver, Ronald P.

Evaluation of a Listening Comprehension Test for Disadvantaged Junior High School Boys. Final Report.

American Institutes for Research, (Washington Office) Silver Spring, Md. Communication Research Program.

Report No-AIR-F-96-6-68-FR

Pub Date Jun 68

Note-24p.

EDRS Price MF-$0.25 HC-$1.30

Descriptors-*Academic Aptitude, Aptitude Tests, *Disadvantaged Youth, Junior High School Students, *Listening Comprehension, Low Income Groups, Males, Middle Class, *Negroes, Standardized Tests, Test Reliability, *Test Validity, Verbal Ability

Identifiers-BoLT, Boys Listening Test, Listening Test, SCAT II, School and College Ability Test Series II, Sequential Tests Educational Progress, STEPLT

This work is a further validation of the findings of an earlier study of the development of a listening test to identify educational potential of disadvantaged Negro junior high school boys. The present study also sought to determine if the experimental boy's listening test (BoLT) is applicable to other ethnic and income level groups. The Bolt, a questionnaire, and two currently used standardized tests of aptitude and listening were administered to 182 low income Negroes, 132 middle income Negroes, 110 low income whites, and 192 middle income whites. Findings show that BoLT is not statistically unique as a measure of educational potential in the low income group. However, the two Negro groups preferred the test while the two white groups did not. Furthermore, there seemed to be no support for the hypothesis that the effect of disadvantagement is associated more with the development of reading proficiency than with verbal proficiency in general. It is concluded that BoLT is an important addition to the area of testing verbal ability and listening comprehension among low income Negro boys. (NH)

ED030687

AIR-F-96-6/68-FR

EVALUATION OF A LISTENING COMPREHENSION TEST FOR DISADVANTAGED JUNIOR HIGH SCHOOL BOYS

Ronald P. Carver

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

FINAL REPORT

Prepared for

College Entrance Examination Board

New York, New York

UD 007 269

American Institutes for Research

Washington Office

Communications Research Program

June 1968

Abstract

Evaluation of a Listening Comprehension Test for Disadvantaged Junior High School Boys

A listening test has been developed to identify educational potential among disadvantaged junior high school boys. In an earlier study, the test, which contains content of interest to this group, was evaluated with the following results and interpretations: (a) the test is reliable and acceptable to this group, (b) the results suggest that the test may be uniquely capable of identifying college potential among disadvantaged students, and (c) the results also suggest that the effect of disadvantagement may be more associated with the development of reading proficiency rather than verbal proficiency in general.

The purpose of the present study was to further validate the findings of the earlier study while extending the evaluation of the test to other ethnic and income level groups. The test, together with a questionnaire and two currently used standardized tests of aptitude and listening, was administered to a large sample of eighth grade boys. Data were analyzed from 182 low-income Negroes, 132 middle-income Negroes, 110 low-income whites, and 191 middle-income whites.

A test-retest study using alternate forms of the test yielded a .78 correlation which rendered further evidence that the test is reliable for low-income Negroes. The high correlation between the test and the standardized listening test (.78) provided concurrent validity for the test as a measure of listening comprehension. This result, together with other results, was interpreted as indicating that the test was not statistically unique as a measure of educational potential among the disadvantaged. The questionnaire responses indicated that the two Negro groups preferred the test over the standardized listening test, while the two white

groups did not prefer the newly developed test. The mean of the middle-income white group was approximately one standard deviation above the mean of the low-income Negroes on all tests including the newly developed listening test. There appeared to be no support for the hypothesis that the effect of disadvantagement is more associated with the development of reading proficiency rather than verbal proficiency in general. Finally, it was concluded that the test is an important addition in the area of testing verbal aptitude and listening comprehension among low-income Negro boys.

Acknowledgments

Many individuals made significant contributions to the research idea, research design, and implementations of the research. The funds and continuing interest necessary to conduct the evaluation were provided by the College Entrance Examination Board. The cooperation, interest, patience and constructive suggestions of Dr. Sam Kendrick of the College Entrance Examination Board were greatly appreciated. Dr. D. B. Orr, formerly Director of AIR's Communication Research Program, was responsible for the research idea in this study and also made suggestions for the research design. Dr. W. R. Graham, formerly of AIR, together with Dr. Orr, was responsible for the original development of the test.

The research was assisted by the Public Schools of the District of Columbia, Dr. William R. Manning, Superintendent. Mr. Norman Nickens, Assistant Superintendent, Model School Division, and Mrs. Barbara Hazel, Assistant, made subjects available in four junior high schools. Mr. John D. Koontz, Assistant Superintendent, Junior and Senior High Schools, made subjects available in three junior high schools.

The Fairfax County Public Schools, Mr. Earl Funderburk, Superintendent, also volunteered to cooperate and Katherine E. Hopper, Supervisor of Guidance, was most helpful in the planning efforts.

The cooperation, assistance, competence, and efforts of the principals and guidance counselors were paramount factors contributing to the success of the testing. Those individuals most directly involved were: Banneker Junior High School, Mr. Elmer Mitchell, Principal, Mrs. Oquino, Department Head, and Mrs. Lightfoot; Gordon Junior High School, Mr. J. Dallas Shirley, Principal, Mrs. Trabue, Department Head, Mrs. Nelson, and Mr. Beasley; Lincoln Junior High School, Mr. Gilbert A. Diggs, Principal, Mr. Louis A. Hartman, Vice Principal, and Mrs. Brown, Department Head; Kramer Junior High School, Miss Alice Morgan, Principal, Mrs.

Gordon, Department Head, Mrs. Mitchell, and Mr. Shackelford; Garnet-Patterson Junior High School, Mrs. Margaret Labat, Principal, Mrs. Tobias, Department Head, and Mrs. Lindia; Alice Deal Junior High School, Mr. Frank A. Stutz, Principal, Mrs. Cochran, Department Head, Mrs. Ontko, and Miss Gross; Shaw Junior High School, Mr. Percy L. Ellis, Jr., Principal, Mrs. Moultrie, Department Head, and Mrs. Holt; Bryant Junior High School, Fairfax County, Virginia, Mr. Nelson W. Coe, Jr., Principal, Mrs. Anne Saunders, Department Head, Mrs. Trimble, and Mr. Hutman.

The recording of the test was accomplished with the assistance of Mr. Warren McDowell, Recording Engineer, Mr. Ernest Brett, Announcer for STEP Listening Test, and Mr. Harold Hairston, Announcer for the experimental listening test. The administration and scoring of the test was accomplished with the assistance of Mrs. Kathleen Styles.

Evaluation of a Listening Comprehension Test for Disadvantaged Junior High School Boys

Background

A listening comprehension test has been developed to identify educational potential among disadvantaged junior high school boys (Graham and Orr, 1966; Orr and Graham, 1968). The content of the test was especially prepared to coincide with the interests of this group. Interests were determined by interviewing boys in the streets of disadvantaged neighborhoods. Stories were then selected to represent the topics of interest indicated in the interviews. The stories, e.g., about spies, detectives, cowboys, were then recorded together with comprehensive multiple choice questions.

The test was evaluated by administering it, together with three standardized aptitude and achievement tests, to a sample of disadvantaged students. The statistical results were interpreted as indicating that the test was reliable, acceptable to the group, and uniquely different from the traditional aptitude and achievement tests. The findings were further interpreted as suggesting (a) that the test was uniquely capable of identifying educational potential among disadvantaged students and (b) that the effect of disadvantagement may be more associated with the development of reading proficiency than with verbal proficiency in general.

The purpose of the present study was to replicate the earlier study while extending the evaluation of the test to other ethnic and income level groups. The testing of other types of groups was necessary in order to investigate the suggested hypotheses concerning the uniqueness of the test and the verbal proficiency of the disadvantaged.

Method

Subjects and Schools. Eighth grade boys in eight schools were tested. Seven schools were in the Washington, D. C. school system and one was in Fairfax County, Virginia, a school within the Metropolitan D. C. area. The District of Columbia Board of Education has designated four junior high schools to be part of a special disadvantaged school district. Most of the eighth grade boys in these four schools were tested. Three other junior highs from the D. C. system and the Fairfax County school were selected so as to provide a sample of different ethnic and income level groups. Although it was intended that all boys in each school be tested, scheduling problems and absentees precluded this possibility. A total of 1084 boys were tested. Subtracted from this total were 121 Ss in a preliminary study, 142 Ss in a reliability study, and 206 Ss who were either foreign students or were students who were absent on one of the test days and, thus, failed to take the complete test battery. Complete data from 615 subjects were analyzed for the main study, including 182 low-income Negroes, 132 middle-income Negroes, 110 low-income whites, and 191 middle-income whites.

Tests and Questionnaire. The School and College Ability Test, Series II (SCAT II), the Sequential Tests of Educational Progress, Listening Test (STEPLT), the experimental boy's listening test (BoLT), and a short questionnaire were administered to each subject.

SCAT II, a new test developed to supplant the earlier version, is a timed test in two parts, verbal analogies (Part I, 20 min.) and arithmetic problems (Part II, 20 min.). The score on Part I (SCAT II Verbal) added to the score on Part II (SCAT II Quant.) yields the score for the test (SCAT II Total).

STEPLT is a traditional standardized listening test. The test is normally read orally by the examiner, but for further standardization

it was tape recorded for the earlier study and the same tape was also used in this study (72 min. on tape).

BoLT is the newly developed listening comprehension test (Orr and Graham, 1968) which includes content designed to be of interest to junior high school boys, e.g., stories about spies, baseball players, and cowboys. For the present study, BoLT was re-recorded in two forms A and B, as indicated by the original study (45 min. for each form). Each form contains 43 multiple choice questions (four alternatives) which are also tape recorded.

The short questionnaire given to each subject included questions related to the occupation and education of parents, tests liked most in the battery, future plans, and attitudes toward school, reading and television.

Income Level. The responses to the questionnaire with regard to occupation and education of parents were used to arbitrarily assign subjects into low and middle-income groups. When the occupations of parents were judged to yield less than $5,000 per annum family income, the subject was placed in the low-income group, and all other subjects were placed in the middle-income group. The nature of the questionnaire responses and the arbitrary judgments required by the experimenter would most likely yield a less than perfect assignment of subjects to income level. However, the schools were selected by the school system as low- or middle-income and the experimenter's judgements on the basis of the student responses was simply a refinement of the first gross categorization. Thus, the errors made in categorization were probably few in number. It should be noted that the income level of low-income whites was probably somewhat higher than low-income Negroes since most of the former were on welfare. Judging from the neighborhoods and the questionnaire responses, the income level of the middle-income Negro group was probably quite comparable to that of the

low-income white group since the middle-income Negroes probably clustered just above $5,000 and the low-income whites just below $5,000.

Testing Procedure. The testing was conducted in the morning on three consecutive days. The guidance department in each school conducted the testing in a manner consistent with their normal testing procedures. The experimenter assisted with the testing in all cases. Some schools required that all students be tested in one large room while others required smaller groups. The tests were always administered in the same order and on consecutive days. SCAT II was given on the first day, STEPLT on the second, and BoLT and the questionnaire on the third. The only exception to this procedure was in one school where the last day of testing was postponed one day due to a teacher's march on Congress.

Preliminary Study. The first administration of the test battery to a group of low-income Negroes included the long form of the BoLT (90 minutes). The impatience of the students toward the end of the test and the frequent laughter elicited by the accent of the announcer prompted the re-recording of the test into two forms by a different Negro announcer.

Reliability Study. In one low-income school both forms of the test were given on two consecutive days. Approximately one-half of the group was administered Form A on the first day and Form B on the second day while the remainder were administered the tests in reverse order.

Results

Table I contains the data for estimating reliability. Inspection of the means reveals no substantial practice effect. The results of the earlier study and the present results confirm that Form B is slightly easier than Form A. Since there appears to be no substantive difference in the means, standard deviations, and correlations due to the order of

administration, the correlation for the total group, .78, can be used as an estimate of the reliability of each form of the test. The .78 alternate form correlation in this study is comparable to the .74 split-half correlation of the earlier study.

TABLE I

Alternate Form Means, Standard Deviations, and Correlations for Two Low-Income Negro Groups

			Form A		Form B		
Group	Order	N	Mean	S.D.	Mean	S.D.	Correlation
I	A,B	51	27.0	8.3	30.4	6.7	.82
II	B,A	67	27.1	6.3	29.5	5.2	.74
Total	Comb.		27.0	7.2	29.9	5.9	.78

Table II contains the intercorrelations among the tests for the total sample and each group separately. For the low-income Negro group the correlations between BoLT (i.e., Form A) and the other test variables are highly similar to the following correlations reported in the earlier study: SCAT Total, .59; SCAT Verbal, .60; SCAT Quantitative, .32; STEPLT, .72. It can be noted that the low-income Negro group in every instance had the lowest correlation of all four groups in the 10 comparison correlations.

The correlations between the STEPLT and BoLT ranged between .65 for the low-income Negroes and .79 for the middle-income Negroes.

TABLE II

Intercorrelations Among the Tests*

		SCAT II			STEPLT	BoLT
	Total	Verbal		Quant.		
SCAT II Total		.91 .92 .92 .94		.90 .91 .92 .92	.63 .76 .68 .80	.52 .64 .64 .66
SCAT II Verbal	.95			.66 .71 .72 .76	.62 .72 .69 .77	.53 .61 .63 .64
SCAT II Quant.	.95	.80			.53 .68 .57 .73	.41 .58 .54 .61
STEPLT	.80	.78		.74		.65 .79 .72 .75
BoLT	.69	.70		.64	.78	
Mean	273.2	267.3		277.3	51.4	32.7
S.D.	17.1	19.1		21.4	12.0	6.6

*Below the diagonal are the correlations, means, and standard deviations for the combined sample, N = 615. Above the diagonal are the correlations for each group according to the matrix:

Low-income Negroes	Middle-income Negroes
Low-income Whites	Middle-income Whites

Figures 1, 2, 3, 4 and 5 contain the means and standard deviations for each of the four income and ethnic groups on each of the five test scores. The low-income Negroes scored lowest on all tests, the middle-income whites scored highest on all test, and the difference between these two groups was always greater than one standard deviation.

The results of the opinionnaire responses are presented in Figure 6. The results are in terms of the precent of the subjects in each group that expressed a preference for BoLT over STEPLT and also the percent that preferred BoLT to SCAT II. It can be noted that all groups preferred BoLT to SCAT II, but only the two Negro groups preferred BoLT to STEPLT.

The intercorrelations among all items on the questionnaire and all of the test scores were computed and inspected for meaning, but none were high enough to be of interest.

In order to obtain an indication of the degree of difference between the two listening tests and the difference between the aptitude test and the listening tests, the data were further analyzed with regard to the number of serious errors of prediction of aptitude in the two listening tests. In order to operationally define serious errors of prediction, additional explanation is necessary. The entire sample of 615 students was used to compute T-scores (Mean = 50, S.D. = 10) for SCAT II (Total), STEPLT, and BoLT. A serious error of prediction was defined as a score on a listening test which was 10 points, one standard deviation, higher than SCAT II. The serious errors of prediction were counted for each group and for each listening test. The number of errors was then converted into percent errors for each group since the number of subjects in each group was not equal. These errors of prediction are termed errors of the first kind and are presented in Table III. Also presented in Table III are errors of the second kind, i.e., errors in which the 10 point difference was negative instead of positive.

FIGURE 1

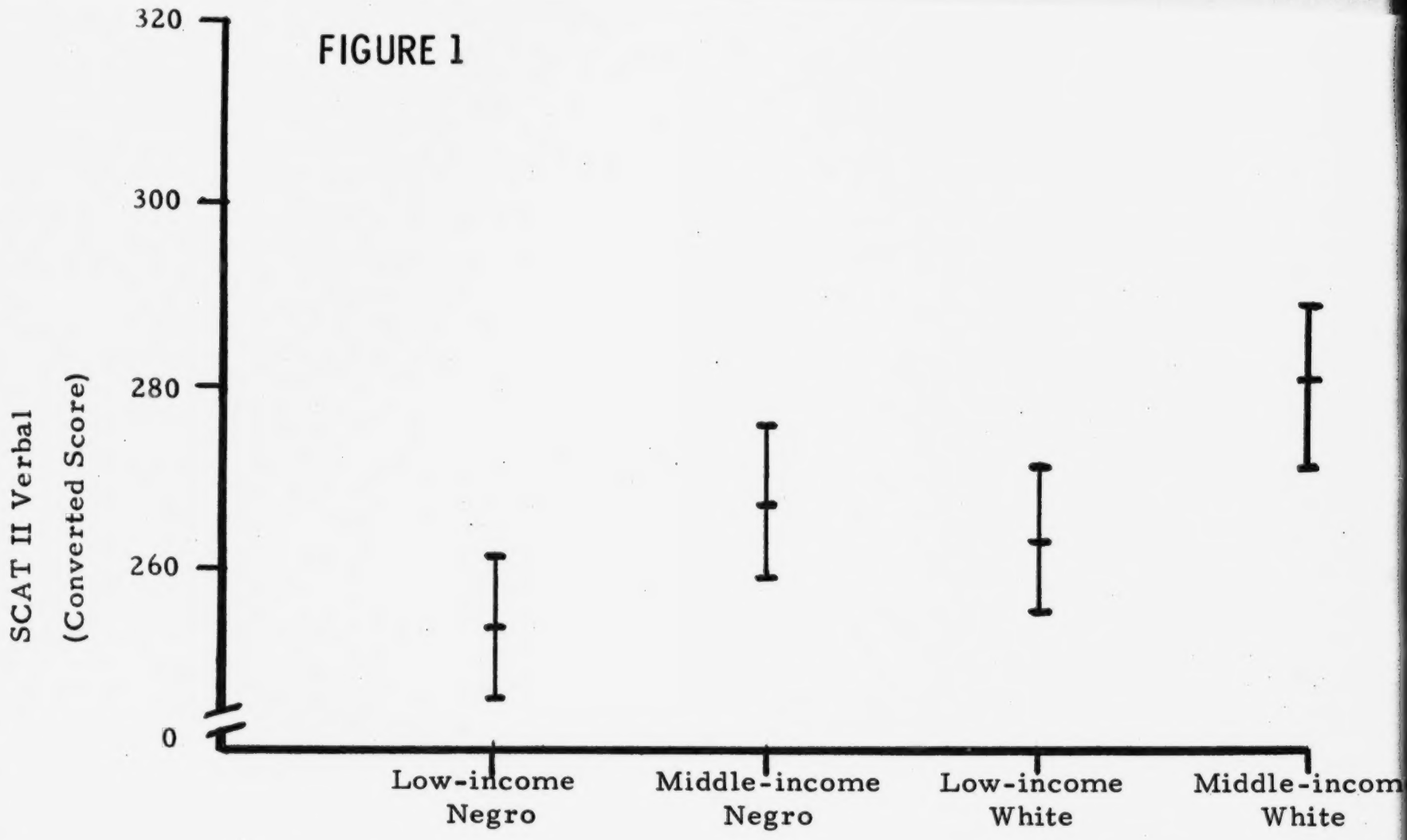

FIGURE 2

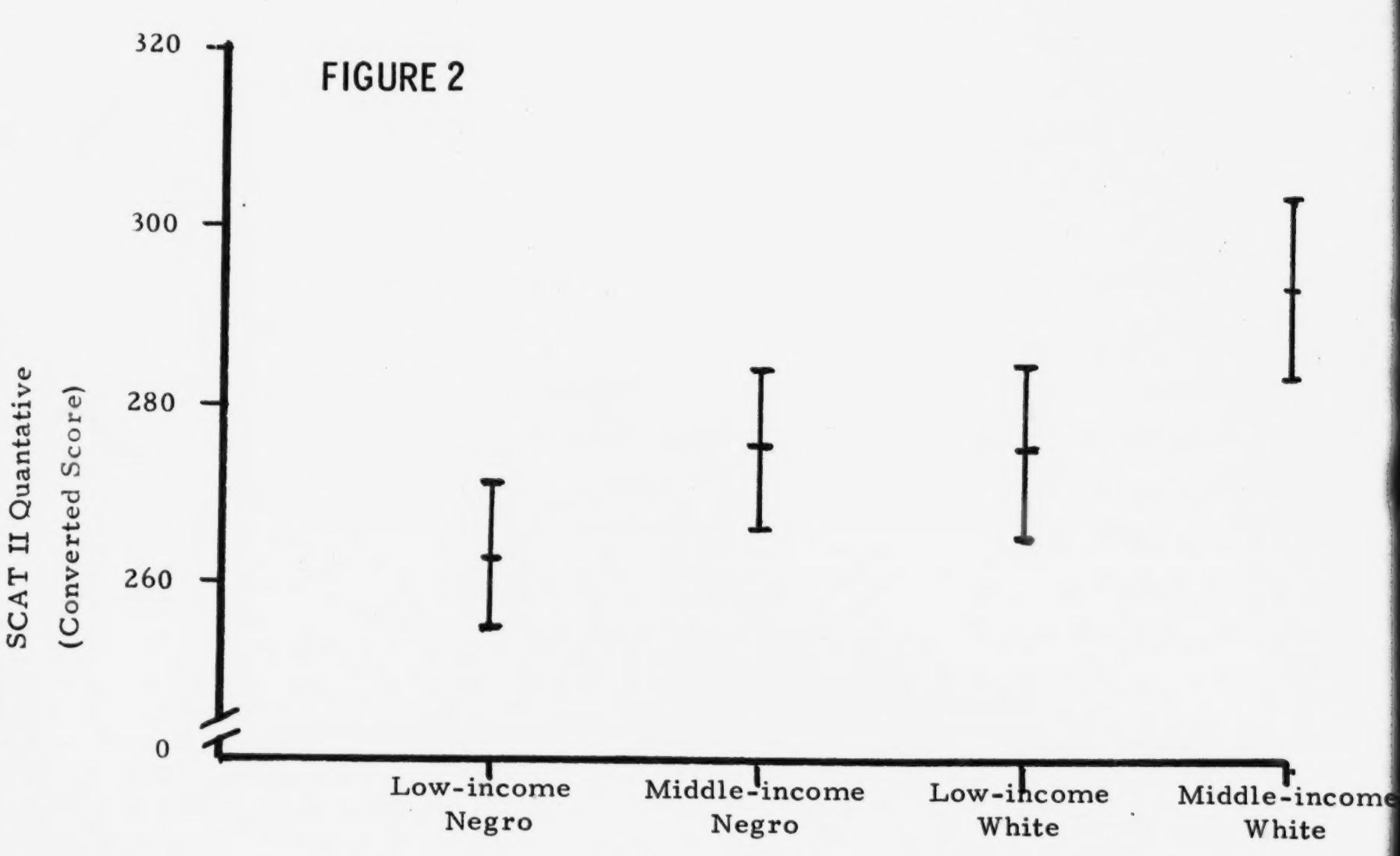

FIGURE 3

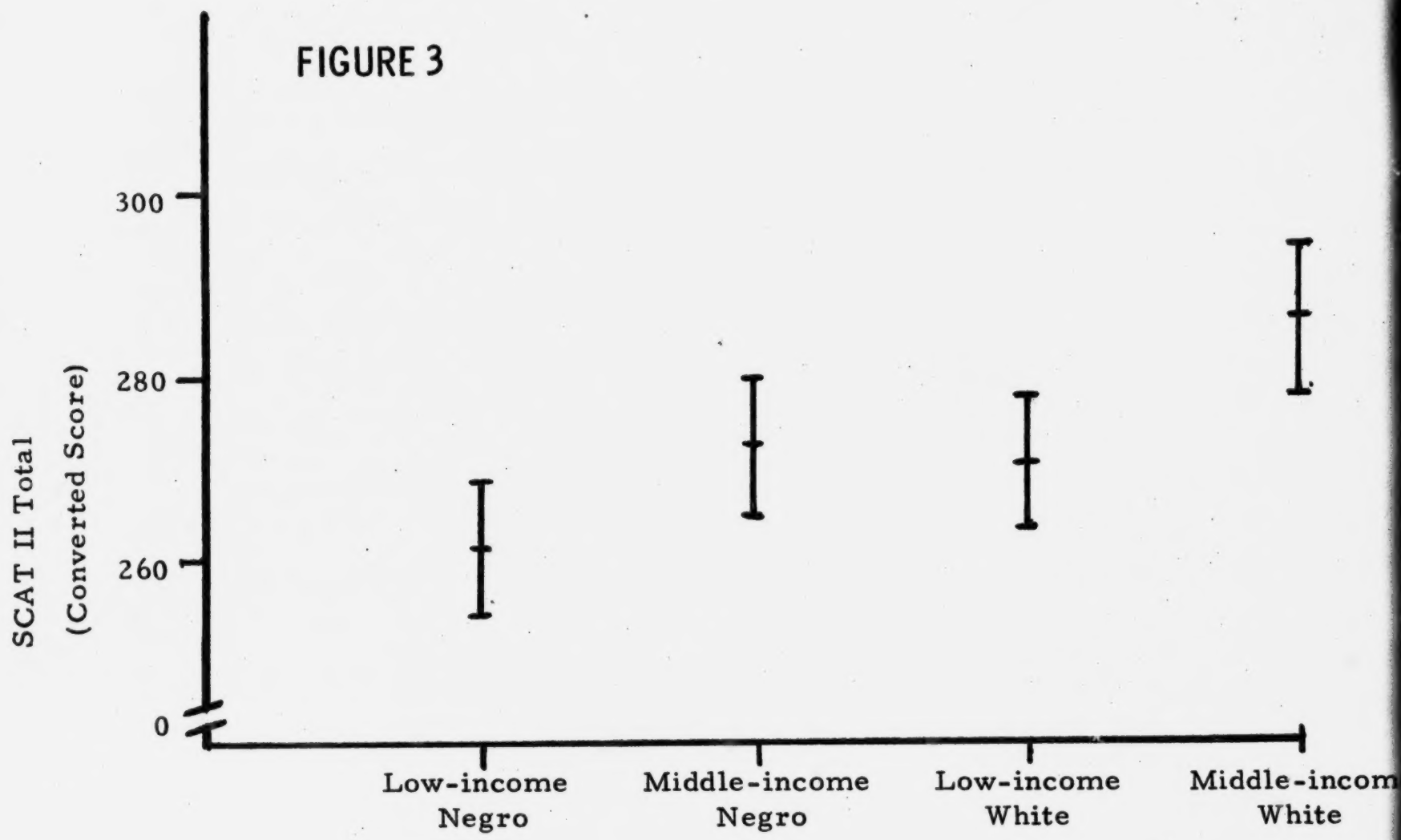

FIGURE 4

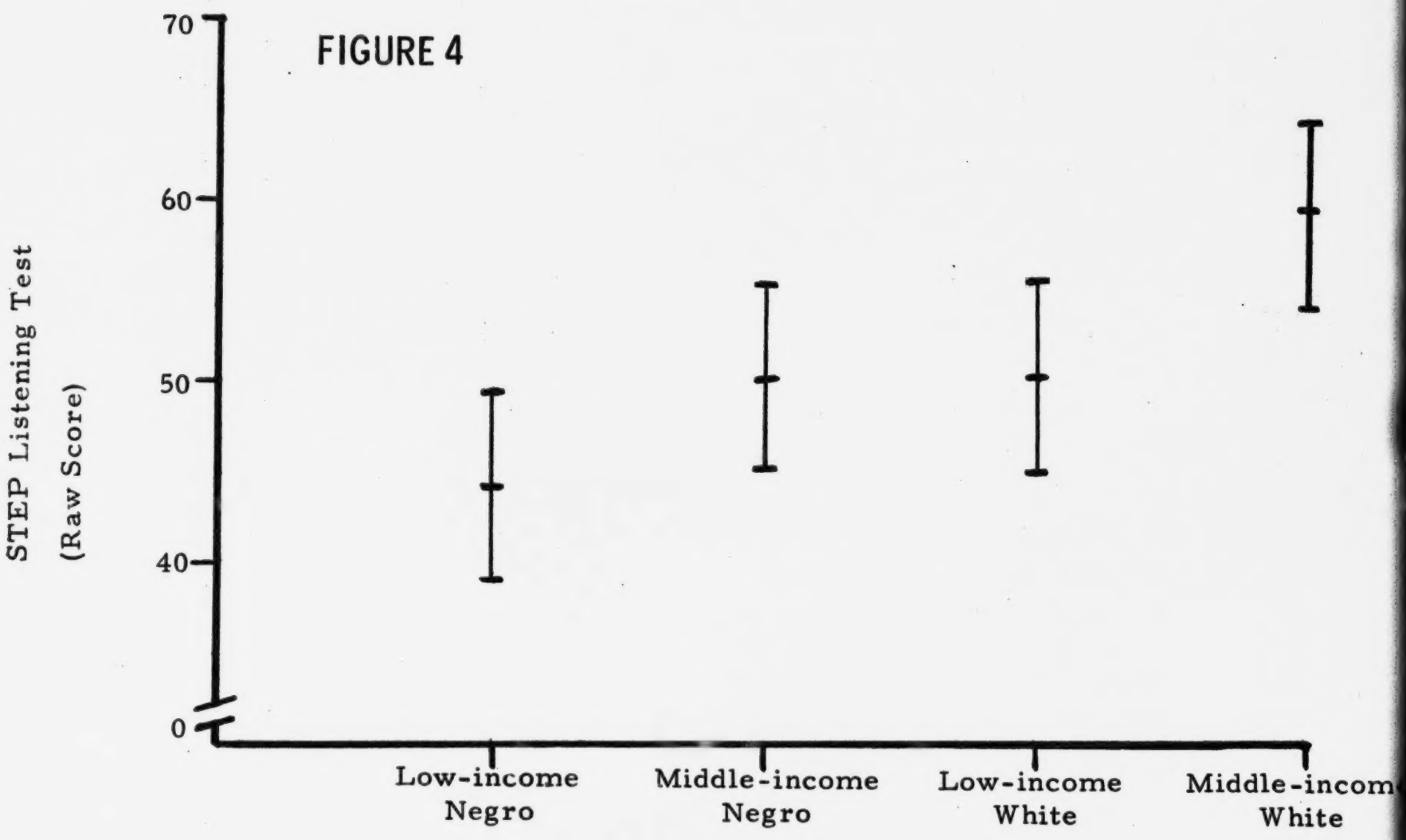

FIGURE 5

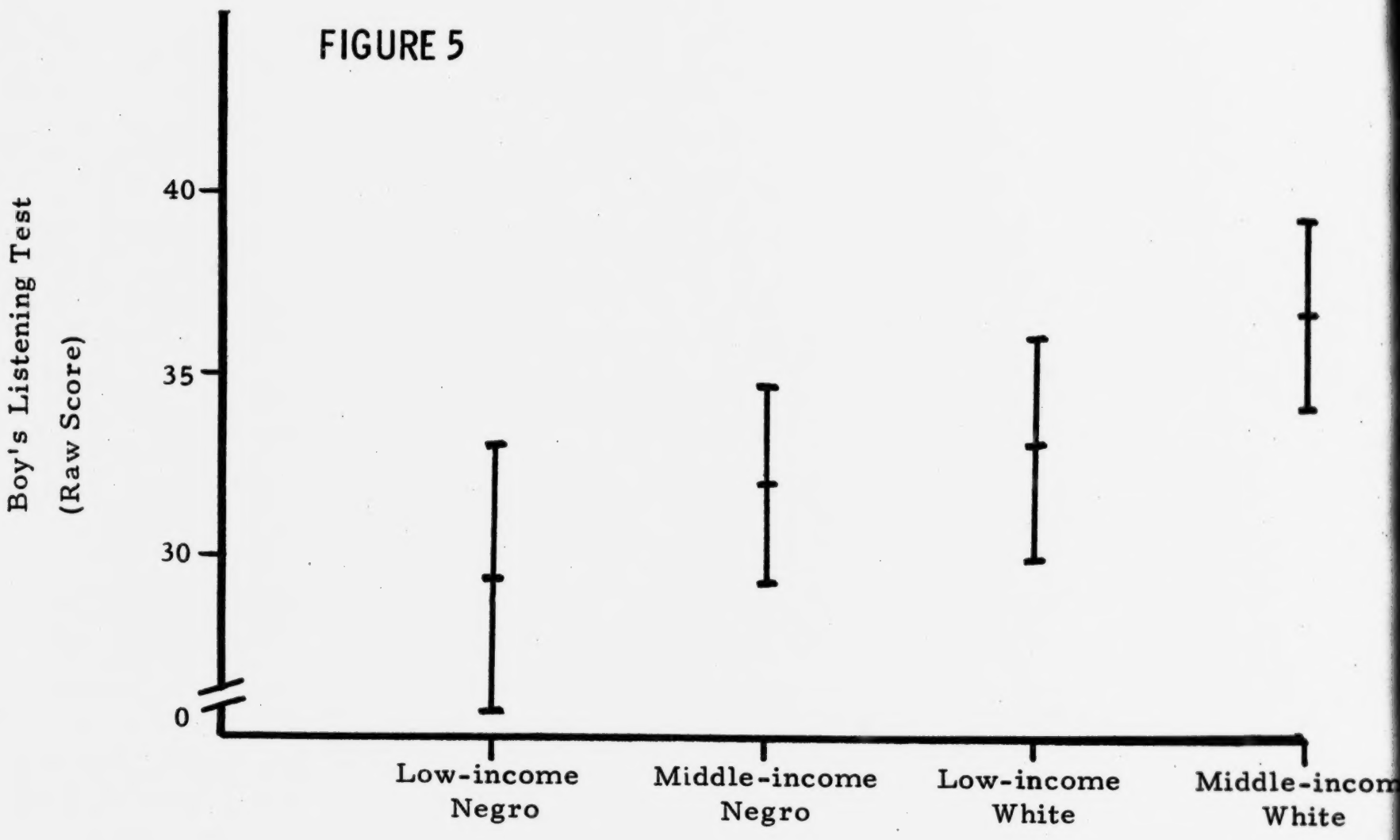

FIGURE 6

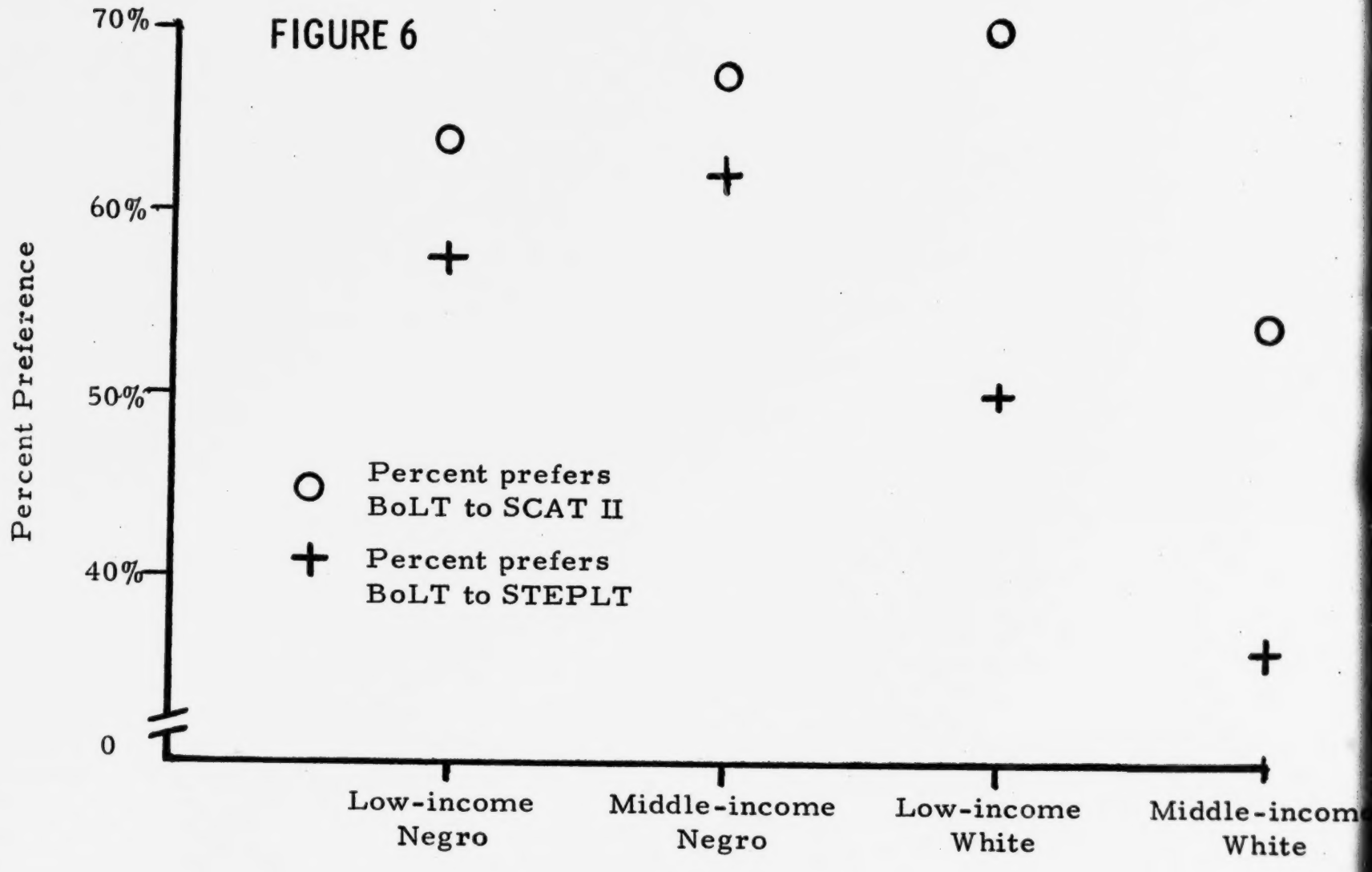

TABLE III

Serious Errors of Prediction*

		First Kind		Second Kind	
		Income		Income	
		Low	Middle	Low	Middle
Ethnic Group	Negro	14.8%	7.6%	9.9%	6.8%
		8.2%	4.5%	5.5%	3.8%
	White	10.0%	6.3%	4.5%	11.0%
		11.8%	2.6%	5.5%	6.8%

*Upper Value: BoLT —SCAT II Total;
Lower Value: STEPLT —SCAT II Total.

It should be noted that for BoLT —SCAT II (Total) there were more errors of the first kind for low-income Negroes than any of the other groups. However, the percent of errors was not high in an absolute sense (14.8%), was not substantially higher than the 8.2% for STEPLT — SCAT II (Total), and different by only 4.9% from the comparable number of errors of the second kind (9.9%).[1]

---

[1]There were no tests of statistical significance due to the following reasons: (a) no known tests were directly applicable, (b) the size of the sample was sufficiently large and the units of measurement were sufficiently meaningful that arbitrary judgments were not considered dangerous, (c) arbitrary judgments concerning size of percents and percent differences were necessary regardless of statistical significance, and (d) Bayesians have appropriately pointed out that statistical procedures have too often been used as symbols of respectability pretending to give the imprimatur of mathmatical logic to the subjective process of empirical inferrence (Edwards, Lindman, & Savage, 1963).

As another way of inspecting the degree to which BoLT and STEPLT are related, the number of common serious errors of prediction were counted and subtracted from the BoLT — SCAT (Total). Common serious errors of prediction are those cases in which an S had a serious error for both STEPLT and BoLT. Table IV contains the percent of errors of prediction for BoLT that are not in common with STEPLT.

TABLE IV

Unique Errors of Prediction for BoLT

		First Kind Income		Second Kind Income	
		Low	Middle	Low	Middle
Ethnic Group	Negro	9.3%	3.0%	8.8%	5.3%
	White	4.6%	4.7%	4.5%	8.4%

Notice that the low-income Negroes again have a larger percent of unique errors of the first kind (9.3%) but this value is not substantially larger than the comparable value for errors of the second kind (8.8%).

Discussion and Conclusions

The reliability of the test for low-income Negroes appears to be adequate and stable since there was little difference between the split-half correlations of the earlier study and the alternate-form correlations in this study. The concurrent validity of the test is quite high, as indicated by the high correlation between the test and the standardized listening test. The test also appears to be an adequate indicator of aptitude since the combined group correlation

with the standardized aptitude test was high. Therefore, it appears that the new listening test is a valid and reliable measure of listening comprehension, and an adequate indicator of aptitude.

The purported high uniqueness of the test for identifying educational potential among the disadvantaged is questionable, however. Carver (1968) has pointed out that the original values used to compute the unique variance of the test are subject to verification and that there is evidence for contending that the test does not have high unique variance with respect to the traditional listening test (article is reproduced in the Appendix). The present results are in accordance with the earlier study and can be similarly interpreted. It is true that the correlation between the STEP Listening Test and the experimental listening test was smaller (.65) for the low-income Negro group than that for any of the other three groups (.79, .72 and .75). However, this result may be attributed to lower reliabilities on the standardized listening test for the low-income Negroes rather than to unique reliable variance. Since the low-income Negro group had the lowest correlation (between the two listening tests) of all four groups in all 10 comparisons, it does suggest that unreliability is a plausible explanation for the lower correlations for the low-income Negro group.

Also the "serious error of prediction" analyses indicate that if the listening test has unique variance, it is not very substantial. The results of the analyses do support, to a certain extent, the uniqueness hypothesis for the listening test in that the low-income Negroes received the largest percent of errors in predicting scholastic aptitude. However, the absolute size of the percent was small (14.8) and when it is compared to several control figures, its magnitude decreases in importance. That is, the comparable values for the other three groups were 7.6, 10.0, and 6.3%, the value comparable to the 14.8% for the STEP Listening Test

was 8.2%, and the errors in prediction in the opposite direction were 9.9%. Furthermore, when errors of prediction possessed in common with the STEP Listening Test are eliminated, the serious errors in prediction for the first kind (9.3%) are approximately equal to the errors of the second kind (8.8%). Thus, it appears that the listening test may have unique variance with respect to other standardized aptitude and listening tests, but the uniqueness is small in magnitude and probably results from less reliable scores in the low-income Negro group.

The test is unique in the sense that it is uniquely preferred by Negroes. Only the two Negro groups preferred the listening test to the STEP Listening Test. The low-income whites were equally split in preference and the middle-income whites preferred the traditional standardized listening test. Although the two white groups did not prefer the test, they both did better on the test than the two Negro groups. From the test score data and the written comments on the back of the questionnaire, it was evident that the test was too easy and a bore to many of the white students.

In the report of the earlier study it was hypothesized that the effect of disadvantagement may be more associated with the development of reading proficiency than with verbal proficiency in general. The results of the current study do not support this hypothesis. The mean of the low-income Negroes was approximately one standard deviation below the mean of the middle-income whites on all measures, not only on verbal and quantitative measures but also on both of the listening tests.

The test was designed for disadvantaged eighth grade boys, and therein lies its assets and limitations. It may not be readily acceptable as a standardized listening test for other groups, such as girls or higher achievement groups. Compared to other tests, it is more likely to

produce valid scores for a disadvantaged Negro group since this group prefers the test, it is at their level of difficulty, and they are thus more likely to be motivated to do their best. Test score norms could be derived from the data collected from the disadvantaged and thus provide meaningful scores for individuals in this group.

The test has other advantages. Since it is a tape recorded test, it is easier to administer and more standardized. Only one hour of testing time is required. The test requires no booklet, and thus can be administered in large numbers very inexpensively. The two parallel forms allow for pre- and post-testing for research purposes. The correlation between the test and the standardized aptitude test was high enough to justify the use of the test as a general measure of aptitude. The test appears to be a valuable addition as a measure of aptitude or listening comprehension among disadvantaged junior high school boys.

In summary, the newly developed listening test (a) is reliable and valid as a listening comprehension test, (b) is preferred by Negro boys as a test of listening comprehension, (c) is not unique as a measure of educational potential among the disadvantaged, (d) does not produce evidence that the effect of disadvantagement may be more associated with the development of reading proficiency rather than verbal proficiency in general, and (e) is an important addition in the area of testing aptitude and listening comprehension among low-income Negro boys.

It is recommended that normative score tables be constructed, a test manual prepared, and the test published for distribution to the public.

References

Carver, R. P. The questionable uniqueness of a newly developed listening test. American Educational Research Journal (In press).

Edwards, W., Lindman, H., & Savage, L. J. Bayesian statistical inference for psychological research. Psychological Review, May, 1963, 70, 193-242.

Graham, W. R., & Orr, D. B. Development of a listening comprehension test to identify college potential among the disadvantaged. Washington: American Institutes for Research, 1966. AIR - E - 48 - 7/66 - FR.

Orr, D. B., & Graham, W. R. Development of a listening comprehension test to identify educational potential among disadvantaged junior high school students. American Educational Research Journal, March, 1968, 5, 167-180.

APPENDIX

THE QUESTIONABLE UNIQUENESS OF A NEWLY DEVELOPED LISTENING TEST

Ronald P. Carver
American Institutes for Research
Washington, D. C.

Orr and Graham (1968) have reported the development of a highly unique listening comprehension test designed to identify educational potential among disadvantaged junior high school students. The high uniqueness of the test was purportedly demonstrated by the finding of a 50 per cent uniqueness coefficient using the following formula given by Flanagan (1962):

where:

$$U_i^2 = r_{ii} - \frac{R^2}{r_{cc}}$$

U_i^2 = uniqueness coefficient for variable i

r_{ii} = reliability coefficient for variable i

R^2 = (multiple) correlation of variable i with the variable(s) in the set

r_{cc} = reliability of the weighted composite of the independent variables

NOTES: (1) The formula given in the Orr and Graham paper incorrectly contains the square root of r_{cc}.

(2) When only two variables are involved, the term R^2/r_{cc} becomes r^2_{ic}/r_{cc} which is the square of the correlation between the two variables when corrected for attenuation (see formula by Thorndike, 1949, p. 107).

The specific values used to compute the 50 per cent uniqueness coefficient were not given. However, by assuming $r_{ii} = .89$ (KR-20 reliability coefficient given for the listening test) and $r_{ic} = .60$ (correlation between the listening test and the aptitude test), the reliability estimate for the aptitude test can be calculated to be .92 when $U^2_i = .50$. The KR-20 reliability given by the test publishers is .95 using 2880 ninth graders in the norm group. For a group of disadvantaged eighth grade boys, the .92 reliability estimate would appear to be extremely high since the group probably scored little better than chance on this particular aptitude test which was designed for 7th, 8th, and 9th grade middle class students. That is, an alternate form reliability coefficient would tend to be low when estimated from a homogeneously low set of scores varying around the chance level. Considering the values used to calculate uniqueness, it appears that the 50 per cent value must represent the upper bound for estimating the uniqueness of the listening test with respect to the aptitude test.

The problem stated by the authors was to determine the uniqueness of the test with respect to traditional aptitude and achievement measures. The authors concluded that the listening test was unique with respect to such tests. Not reported in the paper was the uniqueness of the listening test with respect to one of the achievement measures, the traditional listening test. It seems important to calculate and report this uniqueness when evaluating the newly developed test.

The authors have stated that Form A or Form B can be substituted for the long form of the test with little loss of information. Form A of the test correlated .74 with Form B. Form A correlated .72 with the

traditional listening test. No estimate of the reliability of the traditional listening test is available for this group. However, when the reliability is liberally estimated to be .85 and these three values (r_{ii} = .75; r_{ic} = .72; r_{cc} = .85) are substituted into the uniqueness formula, a coefficient of .13 results.

A listening test may be a better indicator of educational potential among the disadvantaged than traditional aptitude measures, and the newly developed listening test may have certain advantages over a traditional listening test. However, the uniqueness coefficient for the listening test with respect to a traditional aptitude test is probably somewhat lower than 50 per cent and the uniqueness coefficient with respect to a traditional listening test is estimated to be only .13. Therefore, it appears reasonable to question the conclusion that the newly developed listening test is highly unique.

REFERENCES

FLANAGAN, JOHN C., and others. Design for a Study of American Youth. Boston: Houghton-Mifflin, 1962, 240 pp.

ORR, DAVID B., and GRAHAM, WARREN R. "Development of a Listening Comprehension Test to Identify Educational Potential Among Disadvantaged Junior High School Students." American Educational Research Journal 5: 167-180, March 1968.

THORNDIKE, ROBERT L. Personnel Selection. New York: Wiley, 1949, 358 pp.